



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

**Algoritmo primal - dual para el problema de
programación lineal basado en el método de barrera
logarítmica**

TESIS

Para optar el Grado Académico de Magíster en Matemática Pura

AUTOR

Pedro Edgar QUIJANO URBANO

ASESOR

Mg. Edinson Raúl MONTORO ALEGRE

Lima, Perú

2019



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Quijano, P. (2019). *Algoritmo primal - dual para el problema de programación lineal basado en el método de barrera logarítmica*. Tesis para optar grado de Magíster en Matemática Pura. Unidad de Posgrado, Facultad de Ciencias Matemáticas, Universidad Nacional Mayor de San Marcos, Lima, Perú.

HOJA DE METADATOS COMPLEMENTARIOS

CÓDIGO ORCID DEL AUTOR: 0000-0002-8782-7662

CÓDIGO ORCID DEL ASESOR: 0000-0003-1327-2315

DNI: 96271811

GRUPO DE INVESTIGACIÓN:

Ecuación Elíptica Abstracta

INSTITUCIÓN QUE FINANCIA PARCIAL O TOTALMENTE LA INVESTIGACIÓN:

Vicerrectorado de Investigación y Posgrado de la UNMSM.

UBICACIÓN GEOGRÁFICA DONDE SE DESARROLLÓ LA INVESTIGACIÓN. DEBE INCLUIR LOCALIDADES Y COORDENADAS GEOGRÁFICAS

Campus: Ciudad Universitaria (Av. Universitaria s/n – Av. Venezuela cdra. 34, Lima Perú)

Coordenadas: 12⁰03'30" S 77⁰05'00" O

AÑO O RANGO DE AÑOS QUE LA INVESTIGACIÓN ABARCÓ:

2014-2019

ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER

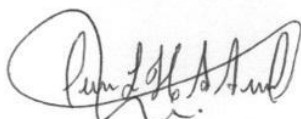
Siendo las, ^{12.00} horas del día viernes dieciocho de octubre del dos mil diecinueve, en el Auditorio de la Facultad de Ciencias Matemáticas, el Jurado Evaluador de Tesis, Presidido por el Dr. Eugenio Cabanillas Lapa e integrado por los siguientes miembros, Dr. Leonardo Alejandro Aguilar (Jurado Evaluador), Mg. Carlos Gilberto Quicaño Barrientos (Jurado Informante); Mg. Willy David Barahona Martínez y el Mg. Edinson Raúl Montoro Alegre como Miembro Asesor, se reunieron para la sustentación de la tesis titulada: «ALGORITMO PRIMAL-DUAL PARA EL PROBLEMA DE PROGRAMACIÓN LINEAL BASADO EN EL MÉTODO DE BARRERA LOGARÍTMICA» presentada por el Bachiller Pedro Edgar Quijano Urbano para optar el Grado Académico de Magíster en Matemática Pura.

Luego de la exposición del graduando, los Miembros del Jurado hicieron las preguntas correspondientes, así como las observaciones e inquietudes acerca del trabajo de tesis, a las cuales el Bachiller Pedro Edgar Quijano Urbano respondió con acierto y solvencia, demostrando pleno conocimiento del tema.

A continuación se realizó la calificación correspondiente, según tabla adjunta, resultando el Bachiller Pedro Edgar Quijano Urbano aprobado con el calificativo de
MUY BUENO (17)

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de **Magíster en Matemática Pura al Bachiller Pedro Edgar Quijano Urbano.**

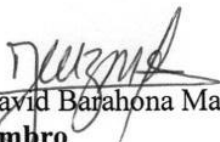
Siendo las ¹³ horas, se levantó la sesión, firmando para constancia la presente Acta.



Dr. Leonardo Alejandro Aguilar
Miembro



Dr. Eugenio Cabanillas Lapa
Presidente



Mg. Willy David Barahona Martínez
Miembro



Mg. Carlos Gilberto Quicaño Barrientos
Miembro



Mg. Edinson Raúl Montoro Alegre
Miembro Asesor

FICHA CATALOGRÁFICA

Pedro QUIJANO

Algoritmo primal - dual para el problema de programación lineal basado en método de barrera logarítmica, (Lima) 2019 VII, 39p. , 29.7cm, (UNMSM, Magister, Matemática Pura, 2019) Tesis, Universidad Nacional Mayor de San Marcos. Facultad de Ciencias Matemáticas.

1. Unidad de Posgrado I. UNMSM/FCM II. Título (Series).

DEDICATORIA

Dedico este trabajo a mi esposa
GABY y a mis hijos DABERT,
GABRIEL y AIRON

AGRADECIMIENTO

Agradesco a mi asesor Edinson MONTORO ALEGRE por la paciencia mostrada durante el tiempo para la elaboración de este trabajo de tesis. También agradezco a mis amigos colegas y familiares por el empuje para la culminación de mi trabajo de tesis. Gracias.

RESUMEN

Algoritmo primal - dual para el problema de programación lineal
basado en método de barrera logarítmica.

Pedro Edgar QUIJANO URBANO

Julio, 2019

Asesor : **Mg. Edinson MONTORO ALEGRE**

Grado obtenido : **Magister en Matemáticas**

En el presente trabajo de tesis, se presenta un método que sigue la trayectoria central para resolver un problema de programación lineal. Las ideas están basadas en el trabajo realizado por Kojima, Mizuno y Yoshise [15] y Monteiro y Adler [18]. El método permite deducir un algoritmo conocido como Algoritmo Primal-Dual de pasos cortos y alcanza una complejidad de orden $O(\sqrt{n} \log \frac{1}{\varepsilon})$ de tiempo [9], debido a que hace uso de una medida de proximidad

$$N_2(\theta) = \{(x, y, s) \in \Omega^0 / \|XSe - \mu e\| \leq \theta\mu\}$$

a la trayectoria central.

Palabras claves: Programación convexa, trayectoria central, método de punto interior, algoritmo Primal - dual.

ABSTRACT

Algorithm primal - dual for the problem of linear programming
based on the logarithmic barrier method

Pedro Edgar QUIJANO URBANO

July, 2019

Assessor : **Mg. Edinson MONTORO ALEGRE**

degree qualification : **Master in Mathematics**

In the present thesis work, we present a method that follow the central trajectory to solve a problem of linear programing. The ideas are based on the work done by KOJIMA, MIZUNO y YOSHISE [15] and MONTEIRO & ADLER [18]. The method allows deducing an algorithm know as algorithm Primal-Dual of shorts steps and reaches a complexity of order $O(\sqrt{n} \log \frac{1}{\varepsilon})$ of time [9], because it makes of a measure of proximity

$$N_2(\theta) = \{(x, y, s) \in \Omega^0 / \|XSe - \mu e\| \leq \theta\mu\}$$

to the central trajectory.

Key words: convex programming, central trajectory, inner point method, primal . dual algorithm.

Índice general

Introducción	1
1. Preliminares	2
1.1. Conjuntos Convexos y Funciones Convexas	2
1.2. Problema de programación lineal estandar (P.P.L. estandar)	4
1.3. Programación convexa	5
1.4. Método de Barrera en Programación Lineal (P.L.)	7
1.5. Método de Newton	11
2. Algoritmo primal-dual siguiendo la trayectoria central para programación lineal	13
2.1. Algoritmo siguiendo trayectoria central	13
2.1.1. Condiciones de Karush-Kuhn-Tucker (KKT)	14
2.2. El método de Newton	17
3. Medidas de proximidad sobre la trayectoria central	21
3.1. Medidas de proximidad	21
4. Algoritmo Primal - Dual de Pasos Cortos	28
4.1. Introducción	28
4.2. Algoritmo Primal - Dual de Pasos Cortos	29
5. Conclusiones	38
Referencias Bibliográficas	39

Introducción

En 1984, con la publicación de un artículo de Karmarkar [12] se da inicio a un nivel de investigación hacia una nueva clase de métodos conocidos como métodos de punto interior y a partir de éstos han emergido diferentes Algoritmos del tipo Primal-Dual como importantes y útiles de esta clase de métodos.

Los métodos Primal-Dual posee excelentes propiedades teóricas y buen desempeño en la práctica.

En el primer capítulo presentamos los conceptos básicos de programación no lineal y programación convexa aplicada al problema de programación lineal. En el capítulo 2 se presenta las condiciones de Karush-Kuhn-Tucker aplicados al problema de Barrera que permite plantear el método Primal-Dual. En el capítulo 3 se presenta la medida de proximidad a la trayectoria central, la cual permitirá plantear el algoritmo. En el capítulo 4 se demuestra los lemas y el teorema que fundamentan la robustez, eficiencia y buen desempeño del algoritmo. Finalmente en el capítulo 5 se presentan las conclusiones.

Capítulo 1

Preliminares

En esta sección presentaremos un resumen de resultados, básicos del análisis convexo, el problema de programación lineal y el método de barrera en programación lineal.

1.1. Conjuntos Convexos y Funciones Convexas

Definición 1.1.1 *Un conjunto $C \subseteq \mathbb{R}^n$ es llamado convexo, si para cualquier par de puntos $x, y \in C$, se tiene:*

$$\lambda x + (1 - \lambda)y \in C \text{ para todo } \lambda \in [0, 1]$$

Definición 1.1.2 *Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo. Una función $f : C \rightarrow \mathbb{R}$ es llamada convexa en C cuando para cualquier $x, y \in C$ y $\lambda \in (0, 1)$, se tiene:*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

La función f es llamada estrictamente convexa en C cuando para cualquier $x, y \in C$ con $x \neq y$, $\lambda \in (0, 1)$, se cumple:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Teorema 1.1.3 *Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ funciones convexas, entonces $f + g$ es una función convexa.*

DEMOSTRACIÓN.- Ver [20], página 56.

□

Proposición 1.1.4 Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo y $f_i : C \longrightarrow \mathbb{R}$, $i = 1, 2, 3, \dots, n$ funciones convexas en C , entonces $\sum_{i=1}^n f_i$ es una función convexa en C .

DEMOSTRACIÓN.- Ver [20], página 56. □

Definición 1.1.5 Sea $C \subset \mathbb{R}^n$ un conjunto convexo, $f : C \longrightarrow \mathbb{R}$ una función convexa. Para el problema

$$(P) \begin{cases} \min f(x) \\ \text{sujeto a} \\ x \in C \end{cases}$$

Diremos que

- (i) $x^* \in C$ es un minimizador global de (P) si $f(x^*) \leq f(x)$; $\forall x \in C$
- (ii) $x^* \in C$ es un minimizador local de (P) si existe $\delta > 0$ / $f(x^*) \leq f(x)$;
 $\forall x \in C \cap B(x^*, \delta)$
- (iii) El conjunto $M = \{x^* \in C / f(x^*) \leq f(x) ; \forall x \in C\}$ es llamado conjunto de minimizadores.

Teorema 1.1.6 (Teorema de minimización convexa) Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo y $f : C \longrightarrow \mathbb{R}$ una función convexa en C . Entonces todo minimizador local del problema

$$\begin{cases} \min f(x) \\ \text{sujeto a} \\ x \in C \end{cases}$$

es minimizador global.

Además, el conjunto de minimizadores es convexo.

Si f es estrictamente convexa entonces no puede haber más de un minimizador global.

DEMOSTRACIÓN.- Ver [20], página 51. □

Definición 1.1.7 Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ que tiene sus segundas derivadas parciales, se define la matriz Hessiana de f como $Hf(x)$ donde

$$Hf(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right] ; \quad \forall x \in \mathbb{R}^n$$

Definición 1.1.8 Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ donde tiene sus primeras y segundas derivadas parciales continuas en $C \subset \mathbb{R}^n$. Diremos que la matriz Hessiana $Hf(x)$

(i) Es definida positiva si $x^t Hf(x)x > 0$; $\forall x \in \mathbb{R}^n$

(ii) Es semidefinida positiva si $x^t Hf(x)x \geq 0$; $\forall x \in \mathbb{R}^n$

Teorema 1.1.9 Sea $C \neq \emptyset$ un conjunto convexo abierto en \mathbb{R}^n y sea $f : C \longrightarrow \mathbb{R}$ una función con segundas derivadas parciales y continuas en C . Entonces f es convexa siempre que, $Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$ es semidefinida positiva para todo $x \in C$.

DEMOSTRACIÓN.- Ver [20], página 54. □

Proposición 1.1.10 Si $Hf(x)$ es definida positiva entonces f es estrictamente convexa.

DEMOSTRACIÓN.- Ver [20] página 54. □

1.2. Problema de programación lineal estandar (P.P.L. estandar)

Un P.P.L. es definido como la minimización o la maximización de una función lineal sujeto a restricciones lineales, así:

$$\left\{ \begin{array}{l} \text{minimizar } C^T x \\ \text{s. a.} \\ Ax = b, \\ x \geq 0, \end{array} \right.$$

$$\text{donde } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}; b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \text{ son vectores donde } x, c \in \mathbb{R}^n \text{ } b \in \mathbb{R}^m \text{ y}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

es una matriz $A = [a_{ij}]_{m \times n}$. Entonces A, b y c son parámetros dados y “ x ” son las variables de decisión.

Observación 1.2.1 $x \geq 0$ denota que $x_i \geq 0$, $i = 1, 2, \dots, n$

1.3. Programación convexa

Sean $f(x), g_1(x), \dots, g_m(x)$ funciones real - valuadas definidas sobre un conjunto $C \subseteq \mathbb{R}^n$. Estamos interesados en el programa

$$(P) \left\{ \begin{array}{l} \text{Minimizar } f(x) \\ \text{s. a.} \\ g_1(x) \leq 0, \\ g_2(x) \leq 0, \\ \vdots \\ g_m(x) \leq 0, \\ \text{donde } x \in C \subseteq \mathbb{R}^n \end{array} \right.$$

Se tiene las siguientes definiciones:

- (i) La función $f(x)$ es llamada función objetivo de (P)
- (ii) Las funciones en las desigualdades $g_1(x) \leq 0$, $g_2(x) \leq 0$, $\dots, g_m(x) \leq 0$ son llamadas restricciones para (P) .
- (iii) Un punto $x \in C$ que satisface todas las restricciones para (P) es llamado punto factible para (P)
- (iv) El conjunto F de todos los puntos factibles para (P) es llamado región factible para (P) .

- (v) Si la región factible para (P) es no vacía, entonces decimos que (P) es consistente.
- (vi) Si existe un punto factible x para (P) tal que $g_i(x) < 0 \quad \forall i = 1, 2, \dots, m$, entonces (P) es llamado superconsistente.
- (vii) Si (P) es un programa consistente y si x^* es un punto factible para (P) tal que $f(x^*) \leq f(x)$ para todo punto factible x para (P) , entonces x^* es una solución para (P) .
- (viii) Decimos que (P) es un programa convexo si la función objetivo $f(x)$, las funciones de restricción $g_i(x)$ y el conjunto $C \subseteq \mathbb{R}^n$ son convexas.
- (ix) A la función $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ con $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}^m$ se llama función Lagrangeana de $f(x)$ y los λ_i , multiplicadores de Lagrange.

Teorema 1.3.1 (*Condiciones de Karush - Kuhn - Tucker (KKT)*) Consideremos el siguiente problema de programación convexa.

$$(P) \begin{cases} \min f(x) \\ \text{s. a.} \\ g_i(x) \leq 0, \quad i = 1, 2, \dots, m \\ x \in C \end{cases}$$

donde $f, g_i : C \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ son funciones convexas, C es convexo sobre \mathbb{R}^n y $x \in \mathbb{R}^n$

Al problema (P) se le asocia una función lagrangeana.

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x), \quad \text{con } \lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}^m$$

Si (P) es superconsistente tal que f, g_i poseen primeras y segundas derivadas parciales continuas sobre C y $x^* \in C^0$. Entonces x^* es solución de (P) si y solo si existe $\lambda^* \in \mathbb{R}^m$ tal que cumple:

- i) $\lambda_i^* \geq 0, \quad i = 1, 2, \dots, m$.
- ii) $\nabla_x L(x^*, \lambda^*) = 0$.
- iii) $\lambda_i^* g_i(x^*) = 0$, para todo $i = 1, 2, \dots, m$.

DEMOSTRACIÓN.- Ver [20] página 185

□

1.4. Método de Barrera en Programación Lineal (P.L.)

Consideremos el siguiente problema de programación lineal en su forma estándar

$$(P) \begin{cases} \text{Minimizar } c^T x \\ \text{sujeto a} \\ Ax = b, \\ x \geq 0. \end{cases}$$

donde $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ y $m < n$.

El conjunto de soluciones factibles del problema estándar y el conjunto de sus puntos interiores asociados al problema son dados respectivamente por:

$$F^P := \{x \in \mathbb{R}^n / Ax = b, x \geq 0\} \text{ es acotado}$$

$$F^0 := \{x \in \mathbb{R}^n / Ax = b, x > 0\} \text{ es no vacío}$$

El respectivo problema dual de (P) será:

$$(D) \begin{cases} \text{Maximizar } b^T y \\ \text{sujeto a} \\ A^T y + z = c, \\ z \geq 0. \end{cases}$$

donde $y \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$. El conjunto de soluciones factibles del problema dual y el conjunto de puntos interiores asociados al problema dual son dados por:

$$T^D = \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n / A^T y + z = c ; \text{ para algun } y \in \mathbb{R}^m, z \geq 0\}$$

$$T^0 = \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n / A^T y + z = c ; \text{ para algun } y \in \mathbb{R}^m, z > 0\} \text{ es no vacío}$$

Ahora definiremos la región primal -dual feasible V^{PD} y su interior V como:

$$V^{PD} = \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n / Ax = b, x \geq 0, A^T y + z = c, z \geq 0\}$$

y

$$V = \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n / Ax = b, x > 0, A^T y + z = c, z > 0\}$$

Asi mismo, es posible al problema (P) convertirlo a un problema con restricciones de igualdad solamente y penalizar las restricciones $x \geq 0$. Esto se logra usando una función barrera, que impida que algunas de las componentes de $x \in \mathbb{R}^n$ lleguen a la frontera ($x_j = 0; j = 1, \dots, n$). Esto se logra al añadir el término $-\ln x_j$ a la función objetivo, este término permite que el valor de la función objetivo se incremente sin límites cuando x_j tiende a 0.

Definición 1.4.1 *la siguiente función:*

$$f(x, \mu) = c^T x - \mu \sum_{j=1}^n \ln x_j, \quad x_j > 0, \quad j = 1, 2, \dots, n$$

es llamada *función barrera logarítmica*, donde μ es un parámetro mayor que cero y $\ln x_j$ denota al logaritmo natural de x_j .

Definición 1.4.2 *La siguiente familia de problemas de programación no lineal:*

$$(P_\mu) \left\{ \begin{array}{l} \text{Minimizar } c^T x - \mu \sum_{j=1}^n \ln x_j, \quad x_j > 0, \quad j = 1, 2, \dots, n \\ \text{sujeeto a} \\ Ax = b, \\ x > 0. \end{array} \right.$$

es conocido como *problema de barrera parametrizado vía parámetro $\mu > 0$* .

El primer término de $f(x, \mu) = c^T x - \mu \sum_{j=1}^n \ln x_j$ mide el valor de la función objetivo del problema lineal estándar y el segundo término funciona como una penalización de los puntos que se aproximan a la frontera de la región factible del problema. El problema (P_μ) define la trayectoria central primal de (P) .

Sea una función

$$\begin{aligned} x : \mathbb{R}^+ &\longrightarrow F^0 \\ \mu &\longrightarrow x(\mu) \end{aligned}$$

Entonces para todo valor del parámetro $\mu > 0$, definimos como $x(\mu)$ a la solución del problema (P_μ) . Además, cuando μ tiende a cero, la curva (o trayectoria) formada por todos los minimizadores es continua en μ y converge a la única solución óptima x^* del problema

(P). Ver [24] página 122.

Para todo esto será necesario considerar algunos supuestos sobre los problemas (P) y (D) .

Imponemos las siguientes hipótesis.

Suposición 1.1

- a) El conjunto $F^0 := \{x \in \mathbb{R}^n / Ax = b, x > 0\}$ es no vacío.
- b) El conjunto $T^\circ := \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n / A^T y + z = c, z > 0\}$ es no vacío.
- c) $Rang(A) = m$.
- d) El problema (P) tiene solución óptima finita.

Bajo esta hipótesis se conocen los siguientes resultados.

Proposición 1.4.3 *Para algún $\mu > 0$, el problema (P_μ) tiene una única solución óptima $x(\mu) \in F^0$.*

DEMOSTRACIÓN.- Vamos a probar que para $\mu > 0$ existe una única solución óptima $x(\mu)$ y además $x(\mu) \in F^0$. Como el problema (P_μ) esta dado

$$\left\{ \begin{array}{l} \text{Min } c^T x + \mu \varphi(x) \\ \text{suje } a \\ Ax = b, \\ x > 0. \end{array} \right.$$

donde $\varphi(x) = -\sum_{j=1}^n \ln x_j$, $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$

$$\begin{aligned} \varphi(x) &= -\ln x_1 - \ln x_2 - \ln x_3 - \dots - \ln x_n \\ &= \varphi_1(x_1) + \varphi_2(x_2) + \varphi_3(x_3) + \dots + \varphi_n(x_n) \end{aligned}$$

entonces

$$\varphi_j(x_j) = -\ln x_j \implies \varphi'_j(x_j) = -\frac{1}{x_j} < 0$$

Además $\varphi''_j(x_j) = \frac{1}{x_j^2} > 0$; $j = 1, 2, \dots, n$

usando teorema (1.1.10) se tiene que $\varphi_j(x_j)$ es estrictamente convexa, $j = 1, 2, \dots, n$

Como $\varphi(x) = \varphi_1(x_1) + \varphi_2(x_2) + \varphi_3(x_3) + \dots + \varphi_n(x_n)$ concluimos usando proposición (1.1.4) que:

$$\varphi(x) = \sum_{j=1}^n \varphi_j(x_j)$$

es estrictamente convexa.

Luego, como $f(x, \mu) = c^T x + \mu \varphi(x)$, con $\mu > 0$ y $\varphi(x)$ es estrictamente convexa y

$c^T x = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$ es convexa, entonces usando el teorema (1.1.3)

$f(x, \mu) = c^T x + \mu \varphi(x)$ es estrictamente convexa.

* Para la existencia de optimo finito el F^0 debe ser cerrado y acotado.

* Si F^0 no es cerrado, puede que f tenga óptimo no finito.

Ejemplo 1.4.4 $f(x) = \frac{1}{\|x\|}$

Por el teorema (1.1.6) (Teorema de minimización convexa) garantiza que como $f(x, \mu)$ es estrictamente convexa, entonces, existe solución única $x(\mu)$.

Ahora por demostrar $x(\mu) \in F^0$.

Como $F^0 := \{x \in \mathbb{R}^n / Ax = b, x > 0\}$ es no vacío, el sistema lineal $Ax = b$ es

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m \end{cases}$$

Luego, cada una de estas ecuaciones lineales son convexas.

Por lo tanto, F^0 es un conjunto convexo.

Luego el problema

$$(P_\mu) \begin{cases} \text{minimizar } f(x, \mu) \\ \text{sujeto a} \\ Ax = b, \\ x > 0. \end{cases}$$

es un programa convexo.

Como existe una única solución optima $x(\mu)$, entonces tiene que satisfacer las condiciones

de KKT (ver Teorema 1.3.1). Entonces existen $y \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$ y la función lagrangeana está dado por:

$$L(x, y, z) = c^T x + \mu \varphi(x) + y(b - Ax) + z(-x)$$

tal que cumple

$$(1.2) \quad \left\{ \begin{array}{ll} (i) & y, z \geq 0 \quad \dots\dots\dots (1) \\ (ii) & \nabla_x L = c + \mu \nabla \varphi(x) - A^t y - z = 0 \\ (iii) & Ax = b, \\ & x^T z = 0, \quad \dots\dots\dots (2) \\ & x > 0. \end{array} \right.$$

Donde $y \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$ son conocidos como multiplicadores de Lagrange.

Por la definición de la función $\varphi(x), (\nabla \varphi(x))_j$ se aproxima a $-\infty$, si x_j se aproxima a cero para algún j . Como el parámetro μ es positivo se tiene que si x_j se aproxima a cero para algún j , entonces $\mu(\nabla \varphi(x))_j$ se aproxima a $-\infty$ y por hipotesis debe tener una solución finita.

Por lo tanto, $x(\mu)$ esta en F^0 es decir en el interior de F^P . Entonces (1.2) se reduce al siguiente sistema:

$$\left\{ \begin{array}{l} c + \mu \nabla \varphi(x) - A^t y = 0 \\ Ax = b, \\ x > 0. \end{array} \right.$$

Pues de (1) y (2) se concluye que $z = 0$ □

1.5. Método de Newton

Sea el sistema de ecuaciones no lineales $f_i(x_1, x_2, \dots, x_n) = 0$; $i = 1, 2, \dots, n$ que se puede expresar simplemente como $F(x) = 0$ donde

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$x \longrightarrow F(x)$$

y $F = (f_1, f_2, \dots, f_n)^T$

El esquema iterativo de Newton esta dado por:

$$x^{k+1} = x^k - \frac{F(x)}{F'(x)} = x^k - [F'(x)]^{-1} F(x^k)$$

donde

$$x^{k+1} = \begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \vdots \\ x_n^{k+1} \end{bmatrix}, \quad x^k = \begin{bmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{bmatrix}, \quad F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad \text{y} \quad F'(x^k) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x^k) & \dots & \frac{\partial f_1}{\partial x_n}(x^k) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^k) & \dots & \frac{\partial f_n}{\partial x_n}(x^k) \end{bmatrix}$$

Asi la sucesión x^k tenderá a la solución del sistema $F(x) = 0$.

Capítulo 2

Algoritmo primal-dual siguiendo la trayectoria central para programación lineal

2.1. Algoritmo siguiendo trayectoria central

Consideremos el problema de programación lineal en su forma estándar como:

$$(P) \left\{ \begin{array}{l} \min c^T x \\ \text{sueto a} \\ Ax = b, \\ x \geq 0. \end{array} \right.$$

donde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ y $m < n$.

El dual de (P) es

$$(D) \left\{ \begin{array}{l} \text{maximizar } b^T y \\ \text{sueto a} \\ A^T y + z = c, \\ z \geq 0. \end{array} \right.$$

donde $y \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$.

En el capítulo 1 hemos definido la región primal-dual factible

$$V^{PD} := \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n / Ax = b, \quad A^T y + z = c, \quad (x, z) \geq 0\}$$

$$V := \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n / Ax = b, \quad A^T y + z = c; \quad (x, z) > 0\} \neq \emptyset$$

Usando la función barrera logarítmica para eliminar las restricciones de desigualdad del problema primal (P), se transforma a un problema del tipo:

$$(P_\mu) \begin{cases} \min c^T x - \mu \sum_{j=1}^n \ln x_j \\ \text{sujeto a} \\ Ax = b. \end{cases}$$

Este es conocido como problema de barrera paramétrico vía parámetro $\mu > 0$.

Observación 2.1.1 *Note que la función barrera logarítmica impide que alguna variable llegue a la frontera ($x_j = 0; j = 1, 2, \dots, n$).*

Usando proposición (1.4.3) garantiza para algún $\mu > 0$, el problema (P_μ) tiene una única solución óptima $x(\mu) \in F^0$, entonces satisface las condiciones de KKT (ver Teorema 1.3.1).

2.1.1. Condiciones de Karush-Kuhn-Tucker (KKT)

Dado el problema de barrera

$$(P_\mu) \begin{cases} \min c^T x + \mu \varphi(x) \\ \text{sujeto a} \\ Ax = b. \end{cases}$$

Se demostró en el desarrollo de la proposición (1.4.3) que $f(x, \mu)$ es estrictamente convexa y $Ax = b$, ecuaciones lineales son convexas esto implica que (P_μ) es un programa convexo.

Luego es equivalente escribir:

$$\left\{ \begin{array}{l} \text{minimizar } f(x, \mu) \\ \text{sujeto a} \\ g_1(x) = b_1 - a_{11}x_1 - a_{12}x_2 - \dots - a_{1n}x_n = 0 \\ g_2(x) = b_2 - a_{21}x_1 - a_{22}x_2 - \dots - a_{2n}x_n = 0 \\ \vdots \\ g_m(x) = b_m - a_{m1}x_1 - a_{m2}x_2 - \dots - a_{mn}x_n = 0 \end{array} \right.$$

Donde f, g_i son funciones con primera y segunda derivada parciales continuas. La función lagrangeana asociada a este problema está dado por,

$$L(x, y) = c^T x + \mu \varphi(x) + y_1 g_1(x) + y_2 g_2(x) + \dots + y_m g_m(x)$$

Aplicando el Teorema 1.3.1 (KKT), (P_μ) es un programa convexo y $f, g_i; i = 1, 2, \dots, m$ funciones con primera y segunda derivadas parciales continuas. Además tiene solución para el problema (P_μ) entonces existen $y_i \in \mathbb{R}, i = 1, 2, \dots, m$ multiplicadores de Lagrange tal que:

i) $y_i \geq 0; i = 1, 2, \dots, m$ por tener restricciones de igualdad. $y = (y_1, y_2, \dots, y_m)^T$.

ii) $\nabla_x L(x, y) = 0 ; x = (x_1, x_2, \dots, x_n)^T$.

$$\nabla_x [c_1 x_1 + c_2 x_2 + \dots + c_n x_n - \mu (\ln x_1 + \ln x_2 + \ln x_3 + \dots + \ln x_n) + y_1 (b_1 - a_{11}x_1 - a_{12}x_2 - \dots - a_{1n}x_n) + y_2 (b_2 - a_{21}x_1 - a_{22}x_2 - \dots - a_{2n}x_n) + \dots + y_m (b_m - a_{m1}x_1 - a_{m2}x_2 - \dots - a_{mn}x_n)] = 0$$

Respecto a x_1 :

$$c_1 - \mu \frac{1}{x_1} = y_1 a_{11} + y_2 a_{21} + \dots + y_m a_{m1}$$

Respecto a x_2 :

$$c_2 - \mu \frac{1}{x_2} = y_1 a_{12} + y_2 a_{22} + \dots + y_m a_{m2}$$

Respecto a x_n :

$$c_n - \mu \frac{1}{x_n} = y_1 a_{1n} + y_2 a_{2n} + \dots + y_m a_{mn}$$

Ordenando

$$\begin{cases} c_1 - \mu \frac{1}{x_1} = y_1 a_{11} + y_2 a_{21} + \dots + y_m a_{m1} \\ c_2 - \mu \frac{1}{x_2} = y_1 a_{12} + y_2 a_{22} + \dots + y_m a_{m2} \\ \vdots \\ c_n - \mu \frac{1}{x_n} = y_1 a_{1n} + y_2 a_{2n} + \dots + y_m a_{mn} \end{cases}$$

En forma matricial tenemos como:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} - \mu \begin{pmatrix} \frac{1}{x_1} & 0 & \dots & 0 \\ 0 & \frac{1}{x_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{x_n} \end{pmatrix} \cdot e = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

$$c - \mu X^{-1}e = A^T y, \quad y \in \mathbb{R}^m, \quad e = (1, 1, \dots, 1)^T$$

$$\text{entonces } A^T y + \mu X^{-1}e = c, \quad y \in \mathbb{R}^m, \quad e \in \mathbb{R}^n$$

iii) $y_i g_i(x) = 0$; $y_i > 0$; $i = 1, 2, \dots, m$.

$$y_i(b_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{in}x_n) = 0 ; \quad i = 1, 2, \dots, m$$

$$\text{entonces } (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) = b_i$$

Se tiene:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

$$\text{entonces } Ax = b$$

Tenemos el siguiente sistema

$$\begin{cases} A^T y + \mu X^{-1}e = c \\ Ax = b. \end{cases}$$

$$\text{Luego, sea } s = \mu X^{-1}e, \text{ multiplicando por } X \text{ se tiene: } Xs = \mu X X^{-1}e = \mu e \quad \delta$$

$$XSe = \mu e, \quad e \in \mathbb{R}^n$$

donde:

$$X = \text{diag}(x_1, x_2, \dots, x_n)$$

$$S = \text{diag}(s_1, s_2, \dots, s_n)$$

$$e = (1, 1, \dots, 1)^T$$

asi obtenemos las condicionales de optimalidad para el problema barrera en la forma primal-dual

$$\begin{cases} A^T y + s - c = 0 \\ Ax - b = 0 \\ XSe - \mu e = 0 \end{cases} \quad (2.1)$$

Llamaremos a (x_μ, y_μ, s_μ) la solución del sistema (2.1) para μ específico. Bajo el supuesto (1.1) con las condiciones (a), (b), (c) y (d) se tiene que, cuando μ tiende a cero, la sucesión $(x(\mu), y(\mu), s(\mu))$ convergen a las soluciones optimas del problema (P) y (D) respectivamente.

si $\mu = 0$ las ecuaciones de (2.1) juntamente con las restricciones $x, s \geq 0$ son exactamente las condiciones de optimalidad de los problemas primal (P) y dual (D). El sistema (2.1) tiene $2n + m$ restricciones y $2n + m$ variables. La utilización del método de Newton para resolver el sistema (2.1), da origen a una familia de métodos de puntos interiores primaldual. Para cada iteración de estos métodos se parte de un punto $(x, y, s) \in V$ tal que x es primal factible, (y, s) es dual factible $y, x, s > 0$. La dirección sobre la cual se camina a partir de este punto se basa en la aplicación del método de Newton a las ecuaciones de (2.1).

2.2. El método de Newton

Para resolver el problema (2.1) por el método de Newton.

Consideremos

$$\begin{aligned} F : \mathbb{R}^{2n+m} &\longrightarrow \mathbb{R}^{2n+m} \\ (x, y, s) &\longmapsto F(x, y, s) \end{aligned}$$

El esquema iterativo Newton está dado por:

$$\tilde{x}_{k+1} = \tilde{x}_k - \frac{F(\tilde{x}_k)}{F'(\tilde{x}_k)} = \tilde{x}_k - [F'(\tilde{x}_k)]^{-1} F(\tilde{x}_k) \quad (2.2)$$

donde F' es el Jacobiano.

Donde

$$F(x, y, s) = \begin{bmatrix} f_1(x, y, s) \\ f_2(x, y, s) \\ f_3(x, y, s) \end{bmatrix} = \begin{bmatrix} A^T y + s - c \\ Ax - b \\ XSe - \mu e \end{bmatrix}$$

y

$$F'(x, y, s) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial s} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial s} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial s} \end{bmatrix} = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}$$

Denotemos los puntos $\hat{\theta}_{k+1} = (x^{k+1}, y^{k+1}, s^{k+1})^T$ y $\hat{\theta}_k = (x^k, y^k, s^k)^T$

En (2.2) se tiene

$$F'(\hat{\theta}_k)(\hat{\theta}_{k+1} - \hat{\theta}_k) = -F(\hat{\theta}_k)$$

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} x^{k+1} - x^k \\ y^{k+1} - y^k \\ s^{k+1} - s^k \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ XSe - \mu e \end{bmatrix}$$

Aquí se trabajara con:

$(x^{k+1} - x^k, y^{k+1} - y^k, s^{k+1} - s^k) = (\Delta x^k, \Delta y^k, \Delta s^k)$ llamada dirección de Newton.

Además los puntos generados están dados como:

$$\begin{aligned} x^{k+1} &= x^k + p\Delta x^k, & k &= 0, 1, 2, \dots, n \\ y^{k+1} &= y^k + p\Delta y^k, & k &= 0, 1, 2, \dots, n \\ s^{k+1} &= s^k + p\Delta s^k, & k &= 0, 1, 2, \dots, n \end{aligned} \tag{2.3}$$

donde, p es longitud de paso, $0 < p \leq 1$.

Además la longitud de paso p elegido para las variables primal y dual es el mismo.

Si $p = 1$, $(x, y, s) \in V$ y $(\hat{x}, \hat{y}, \hat{s})$ es el nuevo punto dado por $(\hat{x} - x, \hat{y} - y, \hat{s} - s) = (\Delta x, \Delta y, \Delta s)$

Luego, tenemos el siguiente sistema:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ XSe - \mu e \end{bmatrix} \tag{2.4}$$

de donde

$$S\Delta x + X\Delta s = -XSe + \mu e$$

$$A\Delta x = 0$$

$$A^T\Delta y + \Delta s = 0$$

Donde $\mu > 0$ es algún parámetro de penalidad.

Es fácil notar que, si $\mu = 0$ el sistema de ecuaciones (2.4) en conjunto con las restricciones de no negatividad para x y s respectivamente representan las condiciones de optimalidad para los problemas primal y dual. La dirección de Newton (2.4) calculado para $\mu = 0$, debe apuntar aproximadamente hacia el optimo de estos problemas x^* y (y^*, s^*) .

Por otro lado, la solución del sistema (2.1) para un $\mu > 0$, es un punto perteneciente a la trayectoria central. En este caso, la dirección de Newton (2.4) debe apuntar aproximadamente hacia la trayectoria central $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$, es decir hacia el interior del octante no negativo. Esto hace que se pueda recorrer un camino a lo largo de esta dirección antes que se alcance la frontera del octante.

En efecto, esto se podrá lograr si consideramos que la dirección dada por (2.4), para algún $\mu > 0$, resulta en un desvío hacia el interior del octante no negativo de la dirección que apunta hacia el óptimo del problema. El objetivo de este desvío es mantenerse alejado de la frontera del octante positivo, de manera que permita un mayor paso a lo largo de la dirección calculado, sin salirse de la región factible del problema.

Si consideramos un punto dado (x, y, s) , y lo reemplazamos en el sistema (2.4), entonces a partir de este sistema se obtendrá una dirección $(\Delta x, \Delta y, \Delta s)$ que apunta a $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$, si elegimos $\mu = \frac{x^t s}{n}$. Se debe observar que la brecha de dualidad asociado a x y (y, s) de los problemas primal y dual es dado por:

$$c^t x - b^t y = c^t x - x^t A^t y = (c - A^t y)^t x = s^t x = x^t s$$

y como $n\mu = x^t s$ entonces la brecha de dualidad puede escribirse como

$$c^t x - b^t y = n\mu$$

Por otro lado la dirección $(\Delta x, \Delta y, \Delta s)$ apunta aproximadamente hacia el punto de la trayectoria central $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ donde $x_i s_i = \mu$ para todo $i = 1, 2, 3, \dots, n$, luego la brecha

de dualidad asociada a este punto central $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ es dado por $x^t(\mu)s(\mu)$ el cual es igual a $n\mu$ es decir es igual a la brecha de dualidad asociada al punto de partida (x, y, s) . A partir de esto se puede desarrollar diferentes tipos de algoritmos.

Capítulo 3

Medidas de proximidad sobre la trayectoria central

3.1. Medidas de proximidad

Consideremos el problema primal

$$(P) \left\{ \begin{array}{l} \min x^t c \\ \text{sujeto a} \\ Ax = b, \\ x \geq 0. \end{array} \right.$$

y su dual

$$(D) \left\{ \begin{array}{l} \max b^t y \\ \text{sujeto a} \\ A^t y + s = c, \\ s \geq 0. \end{array} \right.$$

con $y \in \mathbb{R}^m$, $s \in \mathbb{R}^n$, A matriz $m \times n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$; a " y " se le conoce como variables duales y a " s " como variables duales de olgura.

Por teoría de dualidad se cumple que si (y, s) es viable en el dual y x es viable en el primal, entonces

$$b^t y \leq c^t x$$

y si x^* es el óptimo para el primal y (y^*, s^*) es el optimo en el dual, se cumple $b^t y^* = c^t x^*$
Consideremos el conjunto solución primal-dual como

$$\Omega = \Omega_P \times \Omega_D = \{ (x^*, y^*, s^*) \mid A^t y^* + s^* = c ; Ax^* = b ; (x^*)^t (s^*) = 0 , x^* \geq 0 , s^* \geq 0 \}$$

donde el conjunto solución optimal - primal Ω_P es definido como:

$$\Omega_P = \{ x^* \mid c^t x^* \leq c^t x \ \forall x \in F^P ; Ax^* = b ; x^* \geq 0 \}$$

y el conjunto solución optimal dual Ω_D es definido como

$$\Omega_D = \{ (y^*, s^*) \mid b^t y^* \geq b^t y \ \forall (y, s) \in T^D ; A^t y^* + s^* = c ; s^* \geq 0 \}$$

Definición 3.1.1 (*Curva central*) La trayectoria central \mathcal{C} se define como el conjunto de puntos $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ que resuelve el sistema

$$\left. \begin{aligned} A^t y + s &= c \\ Ax &= b \\ XSe &= \mu e \end{aligned} \right\} \quad (3.1)$$

El método de la trayectoria central sigue a la curva \mathcal{C} en la dirección del decrecimiento de μ , es decir dado un punto (x, y, s) viable en el dual y primal ($x \in F^\circ, (y, s) \in T^\circ$) se elige como parametro $\mu = \frac{x^t s}{n}$, el cual es reemplazado en el sistema de Newton

$$\begin{bmatrix} 0 & A^t & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e - XSe \end{bmatrix} \quad (3.2)$$

La dirección Newton que se obtenga como solución de este sistema apuntará a la solución exacta $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ del sistema (3.1).

Si μ tiende a cero, entonces $\mu = \frac{x^t s}{n} \rightarrow 0$, es decir la brecha de dualidad de los puntos iterados se aproxima a cero.

Esto significa que el método sigue a la curva \mathcal{C} en dirección a una solución en el conjunto Ω .

Bajo esta idea los algoritmos de trayectoria central generan una sucesión de puntos (x^k, y^k, s^k)

con $x^k \in F^\circ, (y^k, s^k) \in T^\circ$ proximos a la curva \mathcal{C} que satisfacen

$$\begin{cases} A^t y + s = c \\ Ax = b \\ x > 0 \\ s > 0. \end{cases}$$

Pero no la ecuación

$$XSe = \mu e$$

esto hace que los puntos iterados se desvien del punto $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ y no esten sobre la trayectoria central \mathcal{C} .

La desviación es medida comparando los productos $x_i s_i$ con el valor $\mu = \frac{x^t s}{n} = \frac{\sum_{i=1}^n x_i s_i}{n}$, asi necesitamos usar la siguiente medida de proximidad entre los puntos (x, y, s) y $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$. Dado $(x, y, s) \in \Omega^\circ$ y $\mu > 0$ la proximidad entre los puntos (x, y, s) y $(x_{(\mu)}, y_{(\mu)}, s_{(\mu)})$ es medida por:

$$d(x, s, \mu) = \frac{1}{\mu} \|XSe - \mu e\| = \frac{1}{\mu} \left\| \begin{pmatrix} x_1 s_1 \\ x_2 s_2 \\ \vdots \\ x_n s_n \end{pmatrix} - \left(\frac{x^t s}{n}\right) e \right\|$$

donde $\|\cdot\|$ puede ser la norma Euclidean $\|\cdot\|_2$ o la norma infinita $\|\cdot\|_\infty$.

Lema 3.1 Si $\frac{1}{\mu} \|XSe - \mu e\| < 1$ entonces $x > 0$, $s > 0$.

DEMOSTRACIÓN.-

En efecto, supongamos que exista $x_i = 0$, $s_i = 0$ para algún i entonces:

$$\frac{1}{\mu} \|XSe - \mu e\|_2 \geq \frac{1}{\mu} |x_i s_i - \mu| = \frac{1}{\mu} |0 - \mu| = 1$$

Lo que es una contradicción, por lo tanto $x > 0$ y $s > 0$. □

Definición 3.1.2 .

(a) Si usamos $\|\cdot\|_2$ y $\theta \in [0, 1 >$ entonces la vecindad sobre la trayectoria $N_2(\theta)$ es definido como: $N_2(\theta) = \{(x, y, s) \in \Omega^\circ / \|XSe - \mu e\|_2 \leq \theta \mu\}$.

(b) Si usamos la norma infinita entonces $\frac{1}{\mu} \|XSe - \mu e\|_{\infty} = \frac{1}{\mu} \max |x_i s_i - \mu|$, entonces:

$$\frac{1}{\mu} |x_i s_i - \mu| \leq \frac{1}{\mu} \|XSe - \mu e\|_{\infty} < \theta$$

$$|x_i s_i - \mu| \leq \theta \mu \quad \forall i = 1, 2, 3, \dots, n$$

entonces podemos definir la vecindad:

$$N_{\infty}(\theta) = \left\{ (x, y, s) \in \Omega^o / \max_i |x_i s_i - \mu| \leq \mu \theta \right\}$$

$$N_{\infty}(\theta) = \{ (x, y, s) \in \Omega^o / \|XSe - \mu e\|_{\infty} \leq \mu \theta \}$$

Vemos que ambas vecindades son muy restrictivas.

(c) Una vecindad menos restrictiva es:

$$N_{-\infty}(\gamma) = \{ (x, y, s) \in \Omega^o / x_i s_i \geq \gamma \mu, \quad \forall i = 1, 2, 3, \dots, n \}$$

donde $\gamma \in < 0, 1 >$

Veamos los gráficos de $N_{-\infty}(\gamma)$ y $N_2(\theta)$

Para $n = 2$

$$N_{-\infty}(\gamma) = \{ (x, y, s) \in \Omega^o / x_i s_i \geq \gamma \mu \quad i = 1, 2 \}$$

donde $\gamma \in < 0, 1 >$ entonces:

$$x_1 s_1 \geq \gamma \left(\frac{x_1 s_1 + x_2 s_2}{2} \right) \quad \dots (1)$$

entonces

$$\begin{aligned} 2x_1 s_1 &\geq \gamma x_1 s_1 + \gamma x_2 s_2 \\ \Rightarrow \left(\frac{2 - \gamma}{\gamma} \right) x_1 s_1 &\geq x_2 s_2 \quad \dots (2) \end{aligned}$$

y

$$x_2 s_2 \geq \gamma \left(\frac{x_1 s_1 + x_2 s_2}{2} \right)$$

de donde se obtiene:

$$\left(\frac{\gamma}{2 - \gamma} \right) x_1 s_1 \leq x_2 s_2 \quad \dots (3)$$

De (2) y (3)

$$\left(\frac{\gamma}{2-\gamma}\right) x_1 s_1 \leq x_2 s_2 \leq \left(\frac{2-\gamma}{\gamma}\right) x_1 s_1$$

Ambos límites, inferior y superior son rectas en \mathbb{R}^2 y como $\lim_{\gamma \rightarrow 0} \left(\frac{\gamma}{2-\gamma}\right) = 0$ y $\lim_{\gamma \rightarrow 0} \left(\frac{2-\gamma}{\gamma}\right) = +\infty$ entonces ellas concurren a los ejes $x_1 s_1$ y $x_2 s_2$ el gráfico es:

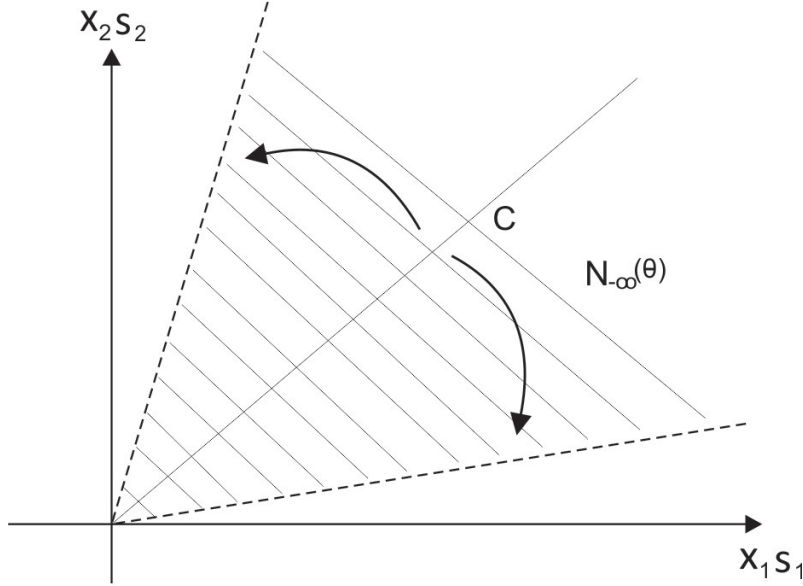


Figura 1

Note que $N_{-\infty}(\gamma)$ tiende a ocupar todo Ω^o para γ muy pequeño.

Ahora analicemos:

$$N_2(\theta) = \{(x, y, s) \in \Omega^o / \|XSe - \mu e\| \leq \theta \mu\}; \quad \theta \in [0, 1 >$$

Tenemos, para $n=2$

$$\begin{aligned} (x_1 s_1 - \mu)^2 + (x_2 s_2 - \mu)^2 &\leq \theta^2 \mu^2 \\ \left(\frac{2x_1 s_1 - x_1 s_1 - x_2 s_2}{2}\right)^2 + \left(\frac{2x_2 s_2 - x_1 s_1 - x_2 s_2}{2}\right)^2 &\leq \theta^2 \mu^2 \\ \left(\frac{x_1 s_1 - x_2 s_2}{2\mu}\right)^2 + \left(\frac{x_2 s_2 - x_1 s_1}{2\mu}\right)^2 &\leq \theta^2 \end{aligned}$$

entonces:

$$\left(\frac{x_1 s_1 - x_2 s_2}{x_2 s_2 + x_1 s_1}\right)^2 \leq \frac{1}{2} \theta^2$$

Luego por la propiedad:

$$-\frac{\sqrt{2}}{2} \theta \leq \frac{x_1 s_1 - x_2 s_2}{x_2 s_2 + x_1 s_1} \leq \frac{\sqrt{2}}{2} \theta$$

donde

$$\underbrace{-\frac{\sqrt{2}}{2} \theta \leq \frac{x_1 s_1 - x_2 s_2}{x_2 s_2 + x_1 s_1} \leq \frac{\sqrt{2}}{2} \theta}_a \quad \wedge \quad -\frac{\sqrt{2}}{2} \theta \leq \overbrace{\frac{x_1 s_1 - x_2 s_2}{x_2 s_2 + x_1 s_1} \leq \frac{\sqrt{2}}{2} \theta}^b$$

De (a):

$$-\frac{\sqrt{2}}{2} \theta x_1 s_1 - \frac{\sqrt{2}}{2} \theta x_2 s_2 + x_2 s_2 \leq x_1 s_1$$

Entonces

$$(1 - \frac{\sqrt{2}}{2} \theta) x_2 s_2 \leq (1 + \frac{\sqrt{2}}{2} \theta) x_1 s_1 \quad \dots\dots(4)$$

$$x_2 s_2 \leq \left(\frac{2 + \sqrt{2} \theta}{2 - \sqrt{2} \theta} \right) x_1 s_1$$

Analogamente de (b) se obtiene:

$$x_2 s_2 \geq \left(\frac{2 - \sqrt{2} \theta}{2 + \sqrt{2} \theta} \right) x_1 s_1 \quad \dots\dots(5)$$

De (4) y de (5) se obtiene:

$$\left(\frac{2 - \sqrt{2} \theta}{2 + \sqrt{2} \theta} \right) x_1 s_1 \leq x_2 s_2 \leq \left(\frac{2 + \sqrt{2} \theta}{2 - \sqrt{2} \theta} \right) x_1 s_1$$

Los límites inferior y superior son rectas, y como:

$$\lim_{\theta \rightarrow 0} \left(\frac{2 - \sqrt{2} \theta}{2 + \sqrt{2} \theta} \right) = 1 = \lim_{\theta \rightarrow 0} \left(\frac{2 + \sqrt{2} \theta}{2 - \sqrt{2} \theta} \right)$$

Entonces, decimos que la banda de esta vecindad $N_2(\theta)$ converge a la curva central por abajo y por arriba.

Observación 3.1.3 .

- (i) Al comparar la figuras 1 y figura 2, se puede concluir que $N_2(\theta)$ es más restrictiva que $N_{-\infty}(\gamma)$

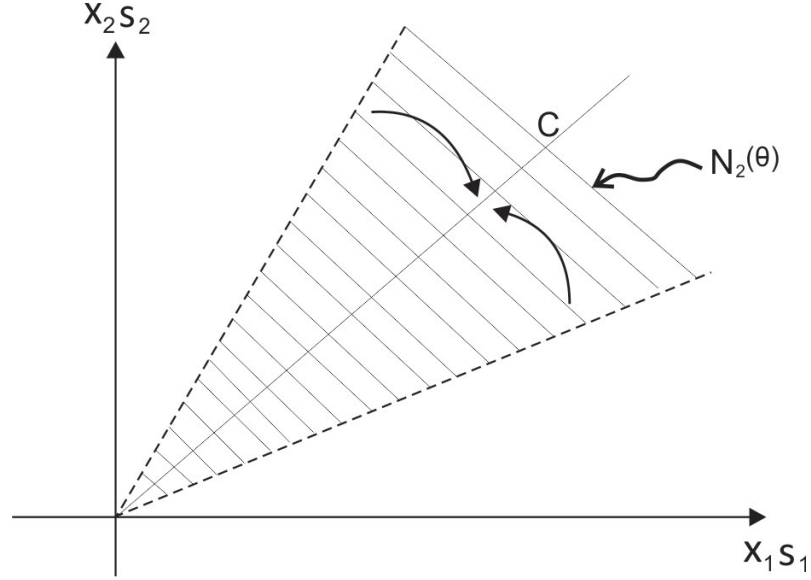


Figura 2

- (ii) Una desventaja de la vecindad $N_2(\theta)$ es que es muy restrictiva. De la definición 3.1.2 (a) se deduce que para que $(x, y, s) \in N_2(\theta)$ se debe cumplir.

$$\sum_{i=1}^n \left(\frac{x_i s_i}{\mu} - 1 \right)^2 \leq \theta^2 < 1$$

Asimismo, si θ esta proximo a 1, la vecindad $N_2(\theta)$ contiene unicamente una pequeña fracción de puntos estrictamente factibles en Ω^0

Capítulo 4

Algoritmo Primal - Dual de Pasos Cortos

4.1. Introducción

Para este tipo de algoritmos, se debe considerar:

- (i) Se inicia con un punto $(x^0, y^0, s^0) \in N_2(\theta)$ con $\alpha_k = 1$, (en el paso Newton); $\tau_k = \tau, \forall k$;
 $0 < \tau < 1$ fijos y la dualidad se va a disminuir por un fracción constante $(1 - \tau)$
- (ii) Veremos que el nuevo punto está en $N_2(\theta)$ es decir si $(x^0, y^0, s^0) \in N_2(\theta)$ entonces todos los iterados siguientes estarán en $N_2(\theta)$.

La idea seguida por los algoritmos de trayectoria central es la siguiente:

Se inicia con un punto $(x^0, y^0, s^0) \in N_2(\theta)$ con $\mu = \frac{(x^0)^t(s^0)}{n}$, entonces el punto (x^0, y^0, s^0) estará próximo al punto de la trayectoria central $(x(\mu), y(\mu), s(\mu))$. Ahora tomamos el valor de τ próxima a 1, entonces tendremos $\tau\mu$ próximo μ , luego esto implicará que el punto (x^0, y^0, s^0) también está próximo a $(x(\tau\mu), y(\tau\mu), s(\tau\mu))$. Esta aproximación garantiza que la dirección de Newton obtenida del sistema

$$\begin{pmatrix} 0 & A^t & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tau\mu e - XSe \end{pmatrix} \dots\dots\dots (I)$$

sea una buena aproximación para la dirección que apunta hacia la trayectoria central y consecuentemente, que el nuevo punto obtenido pertenezca a la "proximidad" de la trayectoria

central definida por $N_2(\theta)$.

Presentamos a continuación uno de los algoritmos de este tipo que fue ideado por Kojima Mizumo e Yoshise.[15]. El cual es llamado Algoritmo de Trayectoria Central de Pasos Cortos por el valor elegido por el τ .

4.2. Algoritmo Primal - Dual de Pasos Cortos

Dado $\varepsilon > 0$, $\theta < 1$, $\tau = 1 - \frac{1}{\sqrt{n}}$, $(x^0, y^0, s^0) \in N_2(\theta)$, se inicia con $k = 0$

Repetir

Hacemos $\tau_k = \tau, \mu_k = \frac{(x^k)^t(s^k)}{n}$

Se calcula la dirección Newton del sistema.

$$\begin{pmatrix} 0 & A^t & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tau_k \mu_k e - X^k S^k e \end{pmatrix} \dots\dots\dots (II)$$

Hacemos: $(x^{k+1}, y^{k+1}, s^{k+1}) = (x^k, y^k, s^k) + (\Delta x^k, \Delta y^k, \Delta s^k)$

$k := k + 1$

Hasta que $\mu_k < \varepsilon$

Lema 4.1 *La dirección $(\Delta x, \Delta y, \Delta s)$ definida por (II) satisface lo siguiente:*

$$(\Delta x)^t(\Delta s) = 0$$

DEMOSTRACIÓN.- Del sistema (II) se obtiene:

$$A^t \Delta y + \Delta s = 0 \text{ entonces } \Delta s = -A^t \Delta y$$

y

$$A \Delta x = 0 \dots\dots\dots (i)$$

Entonces:

$$\begin{aligned} (\Delta x)^t \Delta s &= (\Delta x)^t (-A^t \Delta y) \\ &= -((\Delta x)^t A^t) \Delta y \\ &= -((A \Delta x)^t) \Delta y \\ &= -0^t \Delta y = 0 \end{aligned}$$

□

Lema 4.2 Sea $(x, y, s) \in N_2(\theta)$ y $(\Delta x, \Delta y, \Delta s)$ el paso Newton (al resolver el sistema (II)) y $(\tilde{x}, \tilde{y}, \tilde{s}) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s); \forall \alpha \in [0, 1]$ y $\tilde{\mu} = \frac{\tilde{x}^t \tilde{s}}{n}$ entonces $\tilde{\mu} = [1 - \alpha(1 - \tau)]\mu$

DEMOSTRACIÓN.-

Sabemos

$$\begin{aligned} n\tilde{\mu} &= \tilde{x}^t \tilde{s} = (x + \alpha\Delta x)^t (s + \alpha\Delta s) \\ &= x^t s + \alpha x^t (\Delta s) + \alpha (\Delta x)^t s + \underbrace{\alpha^2 (\Delta x)^t \Delta s}_0 \\ &= x^t s + \alpha [(\Delta x)^t s + x^t \Delta s] \quad \dots (\alpha) \end{aligned}$$

Como $(\Delta x, \Delta y, \Delta s)$ es solución del sistema (II) se cumple:

$$S\Delta x + X\Delta s = \tau\mu e - XSe$$

donde $e = (1, 1, \dots, 1)$

Multiplicamos a ambos lados por e^t :

$$\begin{aligned} e^t [S\Delta x + X\Delta s] &= e^t [\tau\mu e - XSe] \\ e^t S\Delta x + e^t X\Delta s &= \tau\mu n - e^t (XSe) \\ s^t \Delta x + x^t \Delta s &= \tau\mu n - (e^t X)(Se) \\ (\Delta x)^t s + x^t \Delta s &= \tau\mu n - x^t s \quad \dots (\beta) \end{aligned}$$

Pero $\frac{x^t s}{n} = \mu$ y de (α) y (β) :

$$\begin{aligned} n\tilde{\mu} &= x^t s + \alpha(\tau\mu n - n\mu) \\ &= n\mu + \alpha n(\tau - 1)\mu \\ \tilde{\mu} &= (1 - \alpha(1 - \tau))\mu \quad ; \quad \forall \alpha \in [0, 1] \end{aligned}$$

□

Observación 4.2.1 Si $\alpha = 1 \implies \tilde{\mu} = \tau\mu$

y esto ocurre en nuestro algoritmo.

Lema 4.3 Sea $(x, y, s) \in N_2(\theta)$ y $(\Delta x, \Delta y, \Delta s)$ el paso Newton (Del sistema (II)) entonces el nuevo punto. $(\tilde{x}, \tilde{y}, \tilde{s}) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$ con $\tilde{\mu} = \frac{\tilde{x}^t \tilde{s}}{n}$, se cumple:

$$\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq (1 - \alpha)\theta\mu + \alpha^2 \|\Delta X \Delta S e\|$$

DEMOSTRACIÓN.-

$$\begin{aligned} \|\tilde{X} \tilde{S} e - \tilde{\mu} e\| &= \left\| \begin{pmatrix} \tilde{x}_1 & \tilde{s}_1 \\ \tilde{x}_2 & \tilde{s}_2 \\ \vdots & \\ \tilde{x}_n & \tilde{s}_n \end{pmatrix} - \tilde{\mu} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} (x_1 + \alpha\Delta x_1)(s_1 + \alpha\Delta s_1) \\ (x_2 + \alpha\Delta x_2)(s_2 + \alpha\Delta s_2) \\ \vdots \\ (x_n + \alpha\Delta x_n)(s_n + \alpha\Delta s_n) \end{pmatrix} - \tilde{\mu} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} x_1 s_1 + \alpha x_1 \Delta s_1 + \alpha \Delta x_1 s_1 + \alpha^2 \Delta x_1 \Delta s_1 \\ x_2 s_2 + \alpha x_2 \Delta s_2 + \alpha \Delta x_2 s_2 + \alpha^2 \Delta x_2 \Delta s_2 \\ \vdots \\ x_n s_n + \alpha x_n \Delta s_n + \alpha \Delta x_n s_n + \alpha^2 \Delta x_n \Delta s_n \end{pmatrix} - \tilde{\mu} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\| \\ &= \|X S e + \alpha(X \Delta S e + \Delta X S e) + \alpha^2 \Delta X \Delta S e - \tilde{\mu} e\| \quad \dots\dots (*) \end{aligned}$$

Además, por el sistema (II) se tiene:

$$S \Delta x + X \Delta s = \tau \mu e - X S e \quad \dots\dots (**)$$

Pero notemos que:

$$\Delta x = \Delta X e$$

$$\Delta s = \Delta S e$$

Entonces en (**)

$$S\Delta X e + X\Delta S e = \tau\mu e - X S e$$

Reemplazando en (*)

$$\begin{aligned} &= \|X S e + \alpha(\tau\mu e - X S e) + \alpha^2\Delta X\Delta S e - (1 - \alpha(1 - \tau)\mu e)\| \\ &= \|(1 - \alpha)X S e - \mu(1 - \alpha)e + \alpha^2\Delta X\Delta S e\| \\ &= \|(1 - \alpha)(X S e - \mu e) + \alpha^2\Delta X\Delta S e\| \leq (1 - \alpha)\|X S e - \mu e\| + \alpha^2\|\Delta X\Delta S e\| \\ &\leq (1 - \alpha)\theta\mu + \alpha^2\|\Delta X\Delta S e\| \end{aligned}$$

ya que como $(x, y, s) \in N_2(\theta)$ se tiene que $\|X S e - \mu e\| \leq \theta\mu$ □

Lema 4.4 Sea $u, v \in \mathbb{R}^n$ tal que: $u^t v = 0$ entonces

$$\|U V e\| \leq \frac{1}{\sqrt{8}} \|u + v\|^2$$

DEMOSTRACIÓN.- Consideremos la igualdad $u_i v_i = \frac{1}{4}((u_i + v_i)^2 - (u_i - v_i)^2); \forall i = 1, 2, 3, \dots, n$

Entonces:

$$\begin{aligned} \|U V e\|^2 &= \sum_{i=1}^n (u_i v_i)^2 \\ &= \sum_{i=1}^n \frac{1}{16} [(u_i + v_i)^2 - (u_i - v_i)^2]^2 \\ &\leq \frac{1}{16} \sum_{i=1}^n [(u_i + v_i)^4 + (u_i - v_i)^4] \\ &= \frac{1}{16} \left[\sum_{i=1}^n [(u_i + v_i)^2]^2 + \sum_{i=1}^n [(u_i - v_i)^2]^2 \right] \\ &\leq \frac{1}{16} \left[\left(\sum_{i=1}^n (u_i + v_i)^2 \right)^2 + \left(\sum_{i=1}^n (u_i - v_i)^2 \right)^2 \right] \\ &= \frac{1}{16} \left[(\|u + v\|^2)^2 + (\|u - v\|^2)^2 \right] \\ &= \frac{1}{16} [\|u + v\|^4 + \|u - v\|^4] \\ &= \frac{1}{16} (2\|u + v\|^4) \end{aligned}$$

□

Lema 4.5 Si $(x, y, s) \in N_2(\theta)$ y $(\Delta x, \Delta y, \Delta s)$ es el paso Newton entonces

$$\|\Delta X \Delta S e\| \leq \frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)}$$

DEMOSTRACIÓN.- Probaremos a partir del sistema (II).

Como $\Delta x = \Delta X e$; $\Delta s = \Delta S e$, se tiene

$$S \Delta X e + X \Delta S e = \tau \mu e - X S e \quad \dots\dots (i)$$

y usando un cambio de escala. Para esto consideramos a la matriz:

$$D = X^{\frac{1}{2}} S^{-\frac{1}{2}} \quad \dots\dots (ii)$$

$$D^{-1} = S^{\frac{1}{2}} X^{-\frac{1}{2}}$$

Notamos que:

$$\begin{aligned} D^{-1} \Delta x &= S^{\frac{1}{2}} X^{-\frac{1}{2}} \Delta x \\ &= S^{\frac{1}{2}} X^{-\frac{1}{2}} \Delta X e \\ &= S^{-\frac{1}{2}} S^{\frac{1}{2}} S^{\frac{1}{2}} X^{-\frac{1}{2}} \Delta X e \end{aligned}$$

y

$$\begin{aligned} D \Delta s &= X^{\frac{1}{2}} S^{-\frac{1}{2}} \Delta s \\ &= X^{\frac{1}{2}} S^{-\frac{1}{2}} \Delta S e \\ &= X^{-\frac{1}{2}} X^{\frac{1}{2}} X^{\frac{1}{2}} S^{-\frac{1}{2}} \Delta S e \end{aligned}$$

Sumando miembro a miembro

$$\begin{aligned} (D^{-1} \Delta x + D \Delta s) &= S^{-\frac{1}{2}} X^{-\frac{1}{2}} (S \Delta X e) + S^{-\frac{1}{2}} X^{-\frac{1}{2}} (X \Delta S e) \\ &= (S X)^{-\frac{1}{2}} [S \Delta X + X \Delta S] \\ &= (S X)^{-\frac{1}{2}} (\tau \mu e - X S e) \quad \dots\dots (iii) \end{aligned}$$

Por otro lado:

$$\begin{aligned} \Delta X \Delta S &= (D^{-1} \Delta X) (D \Delta S) \\ \|\Delta X \Delta S e\| &= \|(D^{-1} \Delta X) (D \Delta S) e\| \end{aligned}$$

Por el Lema 4.4

$$\begin{aligned}
&\leq \frac{1}{\sqrt{8}} \|D^{-1}\Delta x + D\Delta s\|^2 \\
&= \frac{1}{\sqrt{8}} \|(SX)^{-1/2}(\tau\mu e - XSe)\|^2 \\
&= \frac{1}{\sqrt{8}} \sum_{i=1}^n \frac{(\tau\mu - x_i s_i)^2}{x_i s_i}
\end{aligned}$$

Tambi3n se cumple:

$$\begin{aligned}
&\|XSe - \mu e\| \leq \theta\mu \\
&|x_i s_i - \mu| \leq \|XSe - \mu e\| \leq \theta\mu \\
&-\theta\mu \leq x_i s_i - \mu \leq \theta\mu \\
&\mu - \theta\mu \leq x_i s_i \leq \theta\mu + \mu \\
&\frac{1}{\mu - \theta\mu} \geq \frac{1}{x_i s_i}
\end{aligned}$$

Al aplicarlo:

$$\begin{aligned}
\|\Delta X \Delta Se\| &\leq \frac{1}{\sqrt{8}} \sum_{i=1}^n \frac{(\tau\mu - x_i s_i)^2}{\mu - \theta\mu} \\
&= \frac{1}{\sqrt{8}(1 - \theta)\mu} \|XSe - \tau\mu e\|^2
\end{aligned}$$

Finalmente se prueba:

$$\|XSe - \tau\mu e\|^2 \leq \theta^2 \mu^2 + (1 - \tau)^2 \mu^2 n$$

En efecto:

$$\begin{aligned}
\|XSe - \tau\mu e\|^2 &= \|XSe - \mu e + \mu e - \tau\mu e\|^2 \\
&= \|(XSe - \mu e) + (1 - \tau)\mu e\|^2 \\
&= [(XSe - \mu e) + (1 - \tau)\mu e]^t [(XSe - \mu e) + (1 - \tau)\mu e] \\
&= (XSe - \mu e)^t (XSe - \mu e) + (XSe - \mu e)^t (1 - \tau)\mu e + (1 - \tau)\mu e^t (XSe - \mu e) \\
&\quad + (1 - \tau)\mu e^t (1 - \tau)\mu e
\end{aligned}$$

$$\begin{aligned}
&= \|XSe - \mu e\|^2 + 2(1 - \tau)\mu e^t(XSe - \mu e) + (1 - \tau)^2\mu^2 e^t e \\
&= \|XSe - \mu e\|^2 + 2(1 - \tau)\mu(e^t XSe - \mu e^t e) + (1 - \tau)^2\mu^2 e^t e \\
&= \|XSe - \mu e\|^2 + 2(1 - \tau)\mu \underbrace{\left(\sum_{i=1}^n x_i s_i - n\mu\right)}_0 + (1 - \tau)^2\mu^2 n \\
&= \|XSe - \mu e\|^2 + (1 - \tau)^2\mu^2 n \\
&\leq \theta^2\mu^2 + (1 - \tau)^2\mu^2 n
\end{aligned}$$

Ya que el punto $(x, y, s) \in N_2(\theta)$, esto implicaba que se cumplía $\|XSe - \mu e\| \leq \theta\mu$

luego:

$$\begin{aligned}
\|\Delta X \Delta Se\| &\leq \frac{1}{\sqrt{8}(1 - \theta)\mu}(\theta^2\mu^2 + (1 - \tau)^2\mu^2 n) \\
&\leq \frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)\mu}\mu^2 \\
&= \frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)}\mu
\end{aligned}$$

□

Lema 4.6 Sea $(x, y, s) \in N_2(\theta)$ tal que $(\tilde{x}, \tilde{y}, \tilde{s}) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s) \quad \forall \alpha \in [0, 1]$ donde $(\Delta x, \Delta y, \Delta s)$ es el paso de Newton, $\tilde{\mu} = \frac{\tilde{x}^t \tilde{s}}{n}$ y $\frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)}\alpha \leq \tau\theta$ Entonces para α y τ dado por el algoritmo se cumple:

$$\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq \theta \tilde{\mu}$$

DEMOSTRACIÓN.- Sabemos que por el lema (4.3)

$$\begin{aligned}
\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| &\leq (1 - \alpha)\theta\mu + \alpha^2\|\Delta X \Delta Se\| \\
&\leq (1 - \alpha)\theta\mu + \alpha^2\left(\frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)}\mu\right) \dots\dots\dots (1)
\end{aligned}$$

Como :

$$\frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)}\alpha \leq \tau\theta \quad \dots\dots\dots (2)$$

Luego (2) en (1)

$$\begin{aligned}
&\leq (1 - \alpha)\theta\mu + \alpha\tau\theta\mu \\
&= \theta((1 - \alpha) + \alpha\tau)\mu \\
&= \theta \tilde{\mu}
\end{aligned}$$

Así

$$\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq \theta \tilde{\mu}$$

□

Teorema 4.2.2 Sea $\theta \in]0, 1[$ y $\tau \in]0, 1[$ tal que $\frac{\theta^2 + n(1 - \tau)^2}{\sqrt{8}(1 - \theta)} \alpha \leq \tau\theta$. Si $(x, y, s) \in N_2(\theta)$ y $(\Delta x, \Delta y, \Delta s)$ es el paso de Newton, entonces

$$(\tilde{x}, \tilde{y}, \tilde{s}) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s) \in N_2(\theta); \quad \forall \alpha \in [0, 1]$$

DEMOSTRACIÓN.- El lema (4.6) prueba que: $(\tilde{x}, \tilde{y}, \tilde{s}) \in N_2(\theta)$. Por otro lado, al ser $(\Delta x, \Delta y, \Delta s)$ el paso de Newton se probó que eso implica que:

$$A\tilde{x} = b \quad \text{y} \quad A^t\tilde{y} + \tilde{s} = c$$

Es decir son factibles en el primal y en el dual.

Lo que falta probar es que: $\tilde{x} > 0$ y $\tilde{s} > 0$

En efecto, notamos que: $(\tilde{x}, \tilde{y}, \tilde{s}) \in N_2(\theta)$

Entonces implica:

$$\|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq \theta \tilde{\mu}$$

$$\begin{aligned}
&\left\| \begin{pmatrix} \tilde{x}_1 & \tilde{s}_1 \\ \tilde{x}_2 & \tilde{s}_2 \\ \vdots & \\ \tilde{x}_n & \tilde{s}_n \end{pmatrix} - \tilde{\mu} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\| \leq \theta \tilde{\mu} \\
&|\tilde{x}_j \tilde{s}_j - \tilde{\mu}| \leq \|\tilde{X} \tilde{S} e - \tilde{\mu} e\| \leq \theta \tilde{\mu}
\end{aligned}$$

$$\begin{aligned}
|\tilde{x}_j \tilde{s}_j - \tilde{\mu}| &\leq \theta \tilde{\mu} \\
-\theta \tilde{\mu} &\leq \tilde{x}_j \tilde{s}_j - \tilde{\mu} \leq \theta \tilde{\mu} \\
\tilde{\mu} - \theta \tilde{\mu} &\leq \tilde{x}_j \tilde{s}_j \leq \theta \tilde{\mu} + \tilde{\mu} \\
(1 - \theta) \tilde{\mu} &\leq \tilde{x}_j \tilde{s}_j
\end{aligned}$$

Luego:

$$\tilde{x}_j \tilde{s}_j \geq (1 - \theta) \tilde{\mu} = (1 - \theta)((1 - \alpha) + \alpha \tau) \mu$$

Por otro lado: $0 \leq \alpha \leq 1 \implies -1 \leq -\alpha \leq 0$

Luego: $1 - 1 \leq 1 - \alpha \leq 1 \implies 0 \leq 1 - \alpha \leq 1 \quad \dots (1)$

Por otro lado: $\tau = 1 - \frac{1}{\sqrt{n}} > 0$ y

(i) como $0 < \alpha < 1$

Luego: $0 < \alpha \tau < \tau \implies 0 < \alpha \tau < \tau \quad \dots (2)$

Sumando (1) y (2) se obtiene:

$$0 < (1 - \alpha) + \alpha \tau < 1 + \tau$$

Luego:

$$\tilde{x}_j \tilde{s}_j \geq \underbrace{(1 - \theta)}_{>0} \underbrace{((1 - \alpha) + \alpha \tau)}_{>0} \underbrace{\mu}_{>0} > 0$$

Se tiene probado que :

$$\tilde{x}_j \tilde{s}_j > 0 ; \forall j = 1, 2, \dots, n$$

Lo que implica que: $\tilde{x}_j \neq 0$, $\tilde{s}_j \neq 0$; de donde por Lema 3.1 $\tilde{x}_j > 0$, $\tilde{s}_j > 0$

$\forall j = 1, 2, \dots, n$

(ii) Para $\alpha = 0$ ó $\alpha = 1$ se tiene que $\tilde{x}_j \tilde{s}_j > 0 ; \forall j = 1, 2, \dots, n$ y por el Lema 3.1

$\tilde{x}_j > 0$, $\tilde{s}_j > 0 ; \forall j = 1, 2, \dots, n$

□

Capítulo 5

Conclusiones

Al final del trabajo podemos arribar a las siguientes conclusiones:

- * El algoritmo presentado genera iterados (x^k, y^k, s^k) estrictamente factibles.
- * La desviación de los puntos iterados (x^k, y^k, s^k) de la trayectoria central es medido a traves de

$$\frac{1}{\mu} \|X S e - \mu e\| = \frac{1}{\mu} \left\| \begin{pmatrix} x_1 & s_1 \\ \vdots & \\ x_n & s_n \end{pmatrix} - \left(\frac{x^t s}{n}\right) e \right\|$$

- * Existe una desventaja en la vecindad $N_2(\theta)$, pues es muy restrictiva.

Bibliografía

- [1] J. F. BONNÁNS AND C. C. GONZAGA **Convergence of interior-point algorithms for the monotone linear complementarity problem**, Mathematics of Operations Research, 21 (1996), pp. 1-25.
- [2] V. CHVÁTAL, **Linear Programming**, W. H. Freeman, New York, 1983.
- [3] R. W. COTTLE, J. S. PANG, AND R. E. STONE, **The Linear Complementary Problem**, Academic Press, San Diego, 1992 .
- [4] D. DEN HERTOOG, **Interior Point Approach to Linear, Quadratic, and Convex Programming**, Ph.D. thesis, Technische Universiteit Delft, the Netherlands, September 1992.
- [5] A. V. FIACCO AND G. P. MCCORMICK, **Nonlinear Programming: Sequential Unconstrained Minimization Techniques**, Wiley, New York, 1968. Reprinted by SIAM Publications, 1990.
- [6] K. R. FRISCH **The logarithmic potential method of convex programming**, Technical Report, University Institute of Economics, Oslo, Norway, 1995.
- [7] P. E. GILL, W. MURRAY, M. A. SAUNDERS, J. A. TOMLIN, AND M. H. WRIGHT **On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method**, Mathematical Programming, 36 (1986), pp.183-209.
- [8] C. GONZAGA, **Polynomial affine algorithms for linear programming**, Mathematical Programming, 49 (1990), pp. 7-21.
- [9] C. C. GONZAGA AND M. J. TODD, **An $0(\sqrt{n}L)$ - iteration large-step primal-dual affine algorithm for linear programming**, SIAM Journal on Optimization, 2 (1992), pp. 349-359.
- [10] O. GÜLER, C. ROOS, T. TERLAKY, AND J. P. VIAL, **A survey of the implications of the behavior of the central path for the duality theory of linear programming**, Management Science, 41 (1995), pp. 1922-1934.

- [11] O. GÜLER AND Y. YE, **Convergence behavior of interior-point algorithms**, Management Science, 60 (1993), pp. 215-228.
- [12] N. KARMAKAR, **A new polynomial-time algorithm for linear programming**, Combinatorica, 4 (1984), pp. 373-395.
- [13] L. G. KHACHIYAN, **A polynomial algorithm in linear programming**, Soviet Mathematics Doklady, 20 (1979), pp. 191-194.
- [14] M. KOJIMA, **Basic lemmas in polynomial-time infeasible-interior-point methods for linear programs**, Annals of Operations Research, 62 (1996), pp. 1-28.
- [15] M. KOJIMA, S. MIZUNO, AND A. YOSHISE, **A polynomial-time algorithm for a class of linear complementarity problems** Mathematical Programming, 44 (1989), pp. 1-26.
- [16] K. A. MCSHANE, C. L. MONMA, AND D. F. SHANNO, **An implementation of a primal-dual interior-point method for linear programming**, ORSA Journal on Computing, 1 (1989), pp. 70-83.
- [17] R. D. C. MONTEIRO AND I. ADLER, **Interior path-following primal-dual algorithms. Part I: Linear programming**, Mathematical Programming, 44 (1989), pp. 27-41.
- [18] R. D. C. MONTEIRO, I. ADLER, AND M. G. C. RESENDE, **A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension**, Mathematics of Operations Research, 15 (1990), pp. 191-214.
- [19] Y. E. NESTEROV AND A. S. NEMIROVSKII, **Interior Point Polynomial Methods in Convex Programming**, SIAM Publications, Philadelphia, 1994.
- [20] A.L. PERESSINI, F.E. SULLIVIAN, J.J. UHL, Jr **The Mathematics of Nonlinear Programming**, Mathematics subject classification (1991): 49 - 01 , goC25
- [21] R. A. TAPIA, Y. ZHANG, AND Y. YE, **On the convergence of the iteration sequence in primal-dual interior-point methods**, Mathematical Programming, 68 (1995), pp. 141-154.
- [22] M. J. TODD, **A centered projective algorithm for linear programming**, Mathematics of Operations Research , 15 (1990), pp. 508-529.
- [23] X. XU, **An $O(\sqrt{n}L)$ - iteration large-step infeasible path-following algorithm for linear programming**, Technical Report, Department of Management Sciences, University of Iowa, Iowa City, Ia., 1994.
- [24] LUIS BRICEÑO, CRISTOPHER HERMOSILLA **Optimización no lineal**, Apunte Mat 279